Math 211: Statistics

March 22, 2018

Spring 2018, EXAM 1

__Instructions__: There are five questions in this exam. Please inspect the exam and make sure you have all 5 questions. You may only use your calculator and A4-size **one page** of help sheet. Do all your work on the paper provided.

__Remember:__ ***You must show your work to get proper credit.***

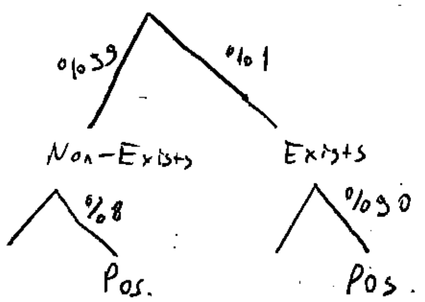
Academic Honesty Code: Koç University Academic Honesty Code stipulates that "copying from others or providing answers or information, written or oral, to others is cheating." By taking this exam, you are assuming full responsibility for observing the Academic Honesty Code.

NAME: Solution

1	/22
2	/20
3	/25
4	/20
5	/23
Total:	/110

10 points bonus

Question 1: (22 Points) **a)** (10 points) A certain genetic defect is known to exist in 1% of the people in the population. As a result of a special test, there is a 90% chance that the test detects the defect. This is called a true positive result. On the other hand, the test gives a false positive result 8% of the time (that is, there is no defect in fact, but the test wrongly concludes that there is). If a person gets a positive test result, what is the probability that s/he actually has the genetic defect?



$$P(E|P_o) = \frac{P(E \cap P_o)}{P(P_o)}$$

$$= \frac{\frac{1}{100} \cdot \frac{90}{100}}{\frac{1}{100} \cdot \frac{9}{100} + \frac{99}{100} \cdot \frac{8}{100}}$$

$$= \frac{90}{882}$$

$$= 0.102$$

b) (6 points) Suppose that 5 people, including you and a friend, line up at random. That is, all 5 of you form a queue at a random order. What is the probability that two people are standing between you and your friend?

5 people can line up in $5!$ different ways.

$\binom{3}{2} \cdot 2!\,2!\,2!$ is the number of ways of two people are standing betwen you and your friend.

So $P = \frac{\binom{3}{2} \cdot 2! \cdot 2!\,2!}{5!} = \frac{24}{120} = \frac{1}{5}$

c) (6 points) A pack of 20 microprocessors has arrived to a store. 4 out of the 20 processors in the pack are actually defective. For quality control, the store manager randomly checks 3 microprocessors. Find the probability that the manager finds at least one defective processor in the check.

Let X be number of defective processors that manager finds

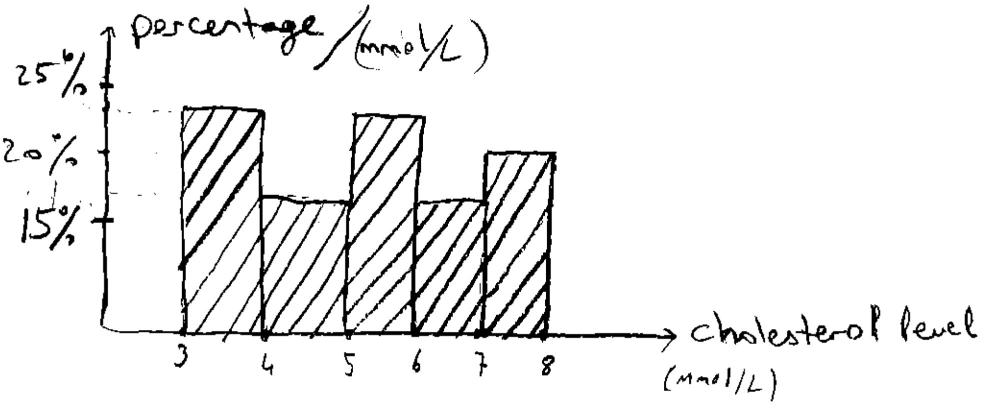
$$P(X \geq 1) = 1 - P(X=0) = 1 - \frac{\binom{16}{3}}{\binom{20}{3}} = 1 - \frac{560}{1140} = 0.509$$

Question 2: (20 Points)

a) (10 points) Draw the density scale histogram of the following data for Serum cholesterol (mmol/L) levels measured on a sample of 25 stroke (=felç) patients.

3.3, 3.4, 3.6, 3.7, 3.8, 3.8, 4.4, 4.5, 4.8, 5.0, 5.4, 5.5, 5.5, 5.6, 5.7, 5.9, 6.1, 6.2, 6.3, 7.0, 7.1, 7.2, 7.6, 7.7, 7.8

Class	$3 < x \leq 4$	$4 < x \leq 5$	$5 < x \leq 6$	$6 < x \leq 7$	$7 < x \leq 8$
frequency	6	4	6	4	5
percentage	$6/25 = 24\%$	16%	24%	16%	20%



b) (10 points) The joint probability distribution of the length X and the width Y (in some specific unit) of parcels in a land for urban development is given in the following table.

y \ x	1	2	3
1	0.10	0.28	0.06
2	0.22	0.22	0.12

Find the distribution of the area $W = XY$ of a randomly selected parcel from this land.

$P_W(1) = P(X=1, Y=1) = 0.1$

$P_W(2) = P(X=1, Y=2) + P(X=2, Y=1) = 0.28 + 0.22 = 0.5$

$P_W(3) = P(X=3, Y=1) = 0.06$

$P_W(4) = P(X=2, Y=2) = 0.22$

$P_W(6) = P(X=3, Y=2) = 0.12$

Question 3: (25 Points) Consider the joint probability density function of X and Y given by

$$f_{X,Y}(x,y) = 2y^3 e^{-xy^2-y^2}, \quad x \geq 0,\ y \geq 0$$

a) (10 points) Find the pdf and the cdf of Y.

$$F_Y(\bar{y}) = P(Y \leq \bar{y}) = \int_0^{\bar{y}} \int_0^{\infty} 2 \cdot y^3 \cdot e^{-y^2(x+1)}\, dx\, dy$$

$$= \int_0^{\bar{y}} \left[-2y\, e^{-y^2(x+1)} \Big|_{x=0}^{x=\infty} \right] dy$$

$$= \int_0^{\bar{y}} 2y\, e^{-y^2}\, dy = 1 - e^{-\bar{y}^2} \Rightarrow f_Y(y) = 2y e^{-y^2}$$

$y > 0.$

$F_Y(y)$

$\downarrow$

b) (10 points) Find $E[XY^2]$.

$$E[XY^2] = \int_0^{\infty}\int_0^{\infty} (xy^2) \cdot 2y^3 e^{-y^2(x+1)}\, dx\, dy$$

$$= \int_0^{\infty} \Bigg[\underbrace{\left(x \cdot (-2y^3 e^{-y^2(x+1)})\right)\Big|_{x=0}^{x=\infty}}_{=0} - \int_0^{\infty} -2y^3 e^{-y^2(x+1)}\, dx \Bigg] dy,$$

by integration by parts, lettings $u = x$, $dv = 2y^5 e^{-y^2(x+1)}$

$$= \int_0^{\infty}\int_0^{\infty} 2y^3 e^{-y^2(x+1)}\, dx\, dy = \int_0^{\infty} \left(-2y e^{-y^2(x+1)}\right)\Big|_{x=0}^{x=\infty} dy = \int_0^{\infty} +2y\, e^{-y^2}\, dy$$

$$= \int_0^{\infty} e^{-z}\, dz$$

$$= -e^{-z}\Big|_{z=0}^{z=\infty}$$

$$= 0 - (-1) = 1$$

c) (5 points) Are X and Y independent? Show.

If X and Y were independent, then $f(x,y) = f_X(x) \cdot f_Y(y)$ would hold.

Equivalently, $f_Y(y) = \dfrac{2y^3 e^{-xy^2-y^2}}{2y\, e^{-y^2}} = y^2 e^{-xy^2}$ would hold, but that depends on x.

$f_Y(y)$ cannot depend on x, so X and Y cannot be independent.

Question 4: (20 Points) Among 3578 small pieces of a special kind of mineral stone obtained from its ore (=maden cevheri), 49 of them are found to have impurity (=another chemical).

If 150 stone pieces are randomly selected and sold to a buyer, what is the probability that the buyer will find at least 3 pieces to contain impurity?

a) (10 points) Use Poisson approximation to Binomial to evaluate the probability in the question.

Let X be the number of pieces that contain impurity

then $X \sim B\left(n=150, p=\frac{49}{3578}=0.0137\right)$

Let $\lambda = n \cdot p = 2.05$, and $Y \sim P(\lambda)$, then

$$P(X \geq 3) \simeq P(Y \geq 3) = 1 - P(Y=0) - P(Y=1) - P(Y=2)$$

$$= 1 - \sum_{k=0}^{2} \frac{e^{-\lambda} \cdot \lambda^{k}}{k!}$$

$$= 0.338$$

b) (10 points) Use the central limit theorem to approximate the probability in the question.

$E[X] = n \cdot p = 2.054$

$Var[X] = n \cdot p(1-p) = 2.026$

$$P(X \geq 3) = P(X \geq 2.5)$$

$$= P\left(\frac{X - E[X]}{\sqrt{Var[X]}} \geq \frac{(2.5) - (2.054)}{1.423}\right)$$

$$\simeq P(Z \geq 0.313), \quad Z \sim N(0,1)$$

$$= 0.3783$$

Question 5: (23 Points) In a physics laboratory, the intensity of light is measured by two independent experimenters from two different sources, respectively. Suppose the intensity measurement X is Normally distributed with mean 25 and standard deviation 3 for both experimenters and the sources.

a) (8 points) Find $P(22 < X < 30)$.

$X \sim N(25, 3^2 = 9)$

$$P(22 < X < 30) = P\left(\frac{22-25}{3} < \frac{X-25}{3} < \frac{30-25}{3}\right)$$

$$\simeq P\left(-1 < z < +\frac{5}{3}\right), \quad z \sim N(0,1)$$

$$= P\left(z < \frac{5}{3}\right) - P(z < -1)$$

$$= 0.9521 - 0.1587$$

$$= 0.7934$$

b) (8 points) If the first experimenter repeats her measurements independently for 70 times, what is the probability that at most 50 of those measurements will be between 22 and 30? (Hint: Use Normal approximation)

Let Y be the number of measurements between 22 and 30, then $Y \sim B(n=70, p=0.7934)$

$E[Y] = np = 55.538$

$Var[Y] = n \cdot p \cdot (1-p) = 11.474$

$$P(Y \leq 50) = P(Y \leq 49.5)$$

$$= P\left(\frac{Y - E[Y]}{\sqrt{Var[Y]}} \leq \frac{(49.5) - (55.538)}{3.387}\right)$$

$$\simeq P(z \leq -1.78)$$

$$(= 1 - 0.9625) = 0.0375$$

c) (7 points) The light is directed to the same target from the two sources. Find the probability that the total intensity on the target is measured to be at least 40.

$X_1, X_2 \sim N(25, 9)$

$X = X_1 + X_2 \sim N(50, 18)$

$$P(X \geq 40) = P\left(\frac{X-50}{\sqrt{18}} \geq \frac{40-50}{\sqrt{18}}\right)$$

$$\simeq P(Z \geq -2.357)$$

$$= 0.99$$